1 Calcolo del valore atteso della varianza campionaria

Ipotesi: $X_1, X_2, \ldots, X_n$ sono un campione casuale di n elementi estratto da un universo $U \sim (\mu,\ \sigma^2)$ (ossia da un universo che presenta qualsiasi forma distributiva, media μ e varianza σ^2). La v.a. media campionaria è definita da $\overline{X} = \sum_{i=1}^{n} X_i/n$ e presenta le seguenti caratteristiche: $E(\overline{X}) = \mu$, $VAR(\overline{X}) = \sigma^2/n$. Lo v.a. varianza campionaria S^2 è definita da

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}$$

Ci chiediamo $E(S^2) =?$, $VAR(S^2) =?$

Per trovare $E(S^2)$ cominciamo a calcolare il valore atteso del numeratore di S^2:

$$\begin{aligned}
E(\sum_{i=1}^{n}\left(X_i - \overline{X})^2\right) &= E\left(\sum_{i=1}^{n}(X_i^2 - 2X_i\overline{X} + \overline{X}^2)\right) \\
&= E\left(\sum_{i=1}^{n} X_i^2 - 2\overline{X}\sum_{i=1}^{n} X_i + n\overline{X}^2\right) \\
&= E\left(\sum_{i=1}^{n} X_i^2 - 2n\overline{X}^2 + n\overline{X}^2\right) \\
&= E\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right) \\
&= \sum_{i=1}^{n} E(X_i^2) - nE(\overline{X}^2)
\end{aligned}$$

Tenendo presente che in una generica variabile casuale Y, $VAR(Y) = E(Y^2) - [E(Y)]^2$ ossia che $E(Y^2) = VAR(Y) + [E(Y)]^2$ e che $\overline{X} \sim (\mu, \sigma^2/n)$ si ottiene:

$$\begin{aligned}
E\left(\sum_{i=1}^{n}(X_i - \overline{X})^2\right) &= \sum_{i=1}^{n}\{VAR(X_i) + [E(X_i)]^2\} - n\left(VAR(\overline{X}) + [E(\overline{X})]^2\right) \\
&= \sum_{i=1}^{n}(\sigma^2 + \mu^2) - n(\frac{\sigma^2}{n} + \mu^2) \\
&= n(\sigma^2 + \mu^2) - n(\frac{\sigma^2}{n} + \mu^2) \\
&= (n-1)\sigma^2
\end{aligned}$$

Di conseguenza:

$$E(S^2) = E\left(\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}\right) = \frac{n-1}{n}\sigma^2.$$

Statistica (A-K) 60 ore
Marco Riani
Università di Parma
http://www.riani.it